United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Applications
Division

SRB Research Report
Number SRB-89-05

March 1989

Forecasting Corn Ear Weight Using Surface Area and Volume Measurements: A Preliminary Report

Fatu G. Bigsby

FORECASTING CORN EAR WEIGHT USING SURFACE AREA AND VOLUME MEASUREMENTS: A PRELIMINARY REPORT, by Fatu Bigsby, Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, March, 1989. Research Report No. SRB-89-05.

ABSTRACT

This study analyzes the forecast performance of models estimated using surface area and volume measurements to predict final corn ear weight. Two models based on research measurements are compared to models estimated using the operational procedures from the Corn Objective Yield Survey. Research and operational models are estimated both within and across years using data from the 1986 and 1987 Michigan Corn Ear Weight Study. Results show that research models based on surface area and volume measurements have mean square errors that are 32 to 52 percent lower than models estimated using the operational procedure. The differences between forecast errors of the research and operational models are statistically significant. The research project was expanded to Michigan and Missouri for 1988 and analysis of that data should provide definitive recommendations concerning the research models.

ACKNOWLEDEMENTS

The author would like to thank staff and field enumerators of the Michigan State Statistical Office for their cooperation, which made this project possible. Support provided by George Hanuschak and Benjamin Klugh is appreciated. The author also thanks Paul Williams for providing assistance in carrying out the project, and Paul Cook for reviewing a draft of this paper.

Washington, D. C. March, 1989

TABLE OF CONTENTS

SUMMARY

The National Agricultural Statistics Service uses the Corn Objective Yield Survey to forecast corn yield, acreage and production. In the Objective Yield Survey, gross yield is defined as the number of corn ears times the average ear weight times the number of sample units in an acre. Net yield, estimated at the state level, is gross yield minus estimated harvest loss. The Corn Ear Weight Study examines one component used to determine yield, specifically, the ear weight.

The purpose of the Corn Ear Weight Study is to develop a more accurate ear weight estimator using surface area and/or volume measurements of the corn ear in addition to ear length measurements made in the operational program. Surface area and volume measurements not only have a potential of producing superior estimates during normal conditions, but especially during drought years. Using ear length alone may not accurately estimate ear weight because of a reduction in the total ear size.

Two research corn ear weight models were estimated using surface area or volume variables. Diameter measurements made one inch from the butt of the cob and two inches from the tip of the cob, and length measurements were used to compute these variables. Research and operational models were estimated, both within and across years, using data from the 1986 and 1987 Michigan Corn Ear Weight study.

Results show that the research models have mean square errors that are 32 to 52 percent lower than those for the operational models. The differences between forecast errors for the research and operational models are statistically significant.

The research project has been expanded to include both Michigan and Missouri for 1988. Since 1988 was a drought year in both states, analysis of the 1988 data, which will be included in a later report, should provide valuable information concerning recommendations to adopt the research model.

Forecasting Corn Ear Weight Using Surface Area and Volume Measurements: A Preliminary Report

Fatu Bigsby[1]

INTRODUCTION

The National Agricultural Statistics Service uses the Corn Objective Yield Survey to forecast corn yield, acreage and production [7,8,10]. In the Objective Yield Survey, gross yield is defined as the number of corn ears times the average ear weight times the number of sample units in an acre. Net yield, estimated at the state level, is gross yield minus estimated harvest loss. The Corn Ear Weight Study examines one component used to determine yield, specifically, the ear weight.

The purpose of the Corn Ear Weight Study is to develop a more accurate ear weight estimator using surface area and/or volume measurements of the corn ear in addition to the length measurements made in the operational program. Surface area and volume measurements not only have a potential of producing superior forecasts during normal conditions, but especially during drought years. Using ear length alone may not accurately forecast ear weight because of a reduction in the total ear size.

The first part of this report gives the background of the study, and the forecasting and validation methodology of the ear weight models. The second part of the report presents results of analyses conducted on the 1986 and 1987 ear weight data for Michigan. In the report, the terms 'current' and 'operational' are interchangeable.

BACKGROUND

The Corn Objective Yield Survey is a probability survey conducted monthly from August to November. Sampling units for the Objective Yield Survey are subsamples of segments selected from the June Enumerative Survey, the major area survey of the National Agricul-

---

[1] The author is a mathematical statistician with the National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C.

tural Statistics Service. Segments are divided into tracts. The number of corn acres planted, or intended to be planted, in the tract determines the probability of selecting a tract. Systematic sampling is the method used to select the tracts. This sampling design selects every kth acre from a list of acres planted, or intended to be planted in the state. The tract selected falls in the kth acre. Each tract contains a sample field which consists of two randomly selected units. The size of the sample unit is fifteen feet long and two rows wide.

A composite of two forecasts ($EW1_{ij}$ and $EW2_{ij}$) produces the monthly ear weight forecast (EW_{ij}). Current measurements used in regression models provide values for $EW1_{ij}$ and $EW2_{ij}$. The models give forecasts for maturity categories 3 through 6, which occur before the crop is harvested.

The regression models are constructed from sample level data from the previous five years. Growing season measurements in the units (cob length) and on the first five ears of corn adjacent to one of the units (kernel row length) provide independent variables for the models. The dependent variable is final ear weight. Both field and laboratory procedures determine the final ear weight when the corn is harvested. The enumerator husks and weighs the corn in row 1 of each unit to obtain a field weight. The laboratory technician determines the moisture content and shelling fraction of the third and fourth ears of each unit. An average ear weight for the sample unit is produced by using the moisture content and shelling fraction to adjust the field weight.

The operational models estimated for maturity categories 3-6 are:

$$\text{Model 1: } EW1_{ij} = \alpha_0+\alpha_1(\text{Kernel Row Length}_{ij})$$

$$\text{Model 2: } EW2_{ij} = \beta_0+\beta_1(\text{Cob Length}_{ij})$$

The method used to obtain a weighted estimate follows:

$$\text{Let } R1 = \Sigma(R^2_{1j} * n_j)/\Sigma n_j \qquad (1)$$

$$R2 = \Sigma(R^2_{2j} * n_j)/\Sigma n_j \qquad (2)$$

where

R^2_{1j} is the coefficient of determination for the kernel row length model and R^2_{2j} is the coefficient of determination for the cob length model estimated for maturity category j. The number of observations in the jth maturity category is n_j. R1 and R2 are calculated by summing across maturity categories 3 through 6.

$$W = R1 + R2$$

$$W1 = R1/W,\ W2 = 1-W1, \text{ and}$$

then the sample ear weight forecast=$(W1*EW1_{ij}) + (W2*EW2_{ij})$

Ronald Steele initiated the Corn Ear Weight Study in 1985 [6]. He modeled the relationship of individual surface area and volume measurements on 78 ears of corn to their final ear weight. The ears were obtained from two purposefully selected fields in Nebraska. Analyses of the Nebraska data showed that models derived from surface area and volume measurements were promising. An earlier study by Vogel [11] showed that a variable computed by multiplying the kernel row length times the midpoint circumference of harvested corn ears had a high correlation with final ear weight.

The Corn Ear Weight Study began formally in 1986 in Michigan. The study differs from the Nebraska analysis in that a probability sample is used to forecast ear weight at the sample level. Steele collected the 1986 data. The author collected the 1987 data and conducted the analyses for both years.

METHODOLOGY

Data Collection

Data collection for the Corn Ear Weight Study continued for two consecutive years during the Corn Objective Yield Survey in Michigan [9]. August 1, September 1, October 1 and November 1 were four survey periods for data collection in 1986. However data collection efforts were reduced to the September 1 and October 1 survey periods in 1987 because the maturity categories used in modeling occur during these months. The maturity categories for the modeling effort were 4 (milk), 5 (dough) and 6 (dent). Maturity category 3 was excluded from the study because of the assumption that size of the corn ear, at this stage of maturity, was not large enough to show a significant relationship to the final ear weight. To preserve consistency between years, analyses in this report only used data collected in September and October during the two years under examination.

Enumerators examined the first five ears of corn next to a designated unit to determine the average maturity of the corn. They measured an ear in the unit if

1) the maturity classification of the sample field was 4, 5, or 6, based on the maturity of the first five ears of corn beyond the specified unit, and

2) the maturity category of the individual ear within the unit was 4, 5 or 6.

The enumerators made three research measurements, over the husk, on ears in row 1 of each unit with a vernier caliper: Length of the kernel row, diameter of the cob two inches from the tip, and diameter of the cob one inch from the butt.

Measuring the length of the kernel row over the husk for the research project differed from the method used in the operational program, where the husk was pulled back. The major objection to the kernel row length measurement over the husk is the difficulty of determining where the row begins and ends. In both the operational and research methods, the enumerator measured a "representative" kernel row, which may have produced a bias in both procedures since the determination of a representative row was subjective.

Appendix D contains examples of questionnnaires (Forms B_1 and C_2) used to collect data for operational program variables which are discussed in this paper. Appendix D also contains the questionnaire used for the Corn Ear Weight Study (Form B_R).

Forecasting Models

Research models for forecasting final ear weight employ surface area, volume or length of the corn ear as independent variables. Surface area and volume computations use a diameter measurement and a length measurement. The diameter measurements are made one inch from the butt of the cob (BD) and two inches from the tip (TD). The length measurements are made for an average kernel row over the husk (KRL), on the kernel row with the husk pulled back (KL), and on the cob over the husk (HL). All measurements take place in row 1 of both units, except for the last two, which are operational program measurements. Cob length over the husk is measured in row 1 of unit 2 and the average kernel row length is measured on the first five ears of corn adjacent to a specified unit during a given month. Sample averages are computed for the measurements.

The definition of a model is

$$Y_i = \beta_o + \beta_1 X_1 + \beta_2 X_2 + e \qquad (3)$$

where,

Y_i=final ear weight for the ith sample field,

X_1, X_2 = surface area, volume, linear or interaction variables (See Figure 1) and,

e_i = the difference between final ear weight and the estimate produced by the model.

The method for estimating the models was Least Squares. For inference purposes, this method assumes that Y_i has a normal distribution with mean $\beta_o+\beta_1X_1+\beta_2X_2$ and variance σ^2; and that e_i has a normal distribution with mean 0 and variance σ^2, Figure 1 gives the variables available for inclusion in a model.

Figure 1. Variables Available for Inclusion in a model

LINEAR VARIABLES:

Kernel row length over the husk(KRL)
Kernel row length with husk pulled back(KL)
Diameter of cob one inch from the butt(BD)
Diameter of cob two inches from the tip(TD)
Length of cob over the husk(HL)

SURFACE AREA VARIABLES:

S1=(HL*BD)
S2=(HL*TD)
S3=(KRL*BD)
S4=(KRL*TD)
S5=(KL*BD)
S6=(KL*TD)

VOLUME VARIABLES:

$V1=(HL*BD^2)/4 * \pi$
$V2=(HL*TD^2)/4 * \pi$
$V3=(KRL*BD^2)/4 * \pi$
$V4=(KRL*TD^2)/4 * \pi$
$V5=(KL*BD^2)/4 * \pi$
$V6=(KL*TD^2)/4 * \pi$

INTERACTION VARIABLE:

BT= Butt diameter * Tip diameter

Although linear and interaction variables are available for inclusion in the research models, the models are forced to contain a surface area or volume variable. The maximum number of independent variables in a model is restricted to two to prevent a high level of multicollinearity.

A Student's t statistic determines if a variable (or its estimated coefficient) is significant and should be retained for further examination. The larger t is, the larger the likelihood that the coefficient is significant. A P value, calculated from t, is the observed significance level of the variable. P values less than or equal to 0.05 determine the significance of variables in this study.

Models are estimated both within and across years. Estimating models both within and across years makes possible the determination of the stability of coefficients across years. This study estimates models across maturity categories while the operational program estimates models within maturity categories. Although there are only two years of data available, models constructed across maturity categories have sufficient data for modeling and evaluation. For comparison purposes, the author assumes that the models respond the same within maturity categories as across categories.

The operational model provides a benchmark for assessing performance of a research model. The author estimated the operational model using two methods. The background section of this paper gave the first method. The exception occurs in calculating R1 and R2 in (1) and (2). R^2's for the kernel row length and husk length models result from fitting the models across all maturity category data.

The second method jointly estimates the relationship of the two operational independent variables to the dependent variable in the same model. Estimation of the operational model using this method produces a more meaningful comparison of the operational and research models because the research model defined in (3) jointly estimates the relationship of the independent variables to the dependent variable. The primary question is: "Is there improvement when the models contain research variables?".

An outlier in the data has a relatively large effect on the magnitude of the estimated coefficients and other regression statistics. Removal of data that are legitimate adversely effects the relevance of the estimated coefficients by decreasing the range over which the coefficients are applicable. Data identified as outliers are not deleted unless confirmed that the data are invalid. New estimates for the model coefficients occur only after deletion of one or more data points which are considered invalid. Appendix A gives the procedure for outlier detection.

Data for refusals, lost samples and inaccessibles are treated as missing values. The models are not modified to account for missing values.

Validation

A model's estimation statistics show its potential for forecasting, but this potential needs independent verification. Validation procedures provide this type of verification.

In this research, validation serves two purposes. First, validation determines if estimated coefficients remain stable across years. A model should have stable coefficients to allow better interpretation of the model coefficients, permit combination of data across years to estimate the parameters, and make the model developed in one year applicable to the next. Secondly, validation determines the accuracy with which the models predict ear weight. A model's ability to forecast is a primary indicator of its utility.

Stability of Coefficients

In validating a model to determine if its coefficients are stable, the data is divided by time. The model is estimated by using data representing each period and by using the combined data. The combined data includes all of the observations for both periods. In this study this means that the model is estimated using the data for 1986, the data for 1987 and the data for both years. Appendix B provides a statistical test that evaluates the stability of estimated coefficients by determining if coefficients estimated for the first year are the same as coefficients estimated for the second year.

Accuracy In Forecasting

In validating a model for its ability to forecast accurately, data independent of the estimation data is used. Validation gives an indication of the forecasting ability of a model by determining how the model performs using data other than that used for its estimation. Figure 2 shows the data sets used to accomplish this.

Figure 2. Validation for Forecasting Accuracy: Data Used

Model Built	Estimation Data	Validation Data
Within Years	1986, n=130	1987, n=131
	1987, n=131	1986, n=130
Across Years	1986 and 1987; a random sample of 1/2 of Sept. Data and all of Oct.'s data; n=186.	1986 and 1987; a random sample of 1/2 of Sept. Data. n=76: 36 obs. for 1986 and 40 for 1987

Note that in validating the model for forecasting accuracy, the combined data does not include all of the observations for 1986 and 1987 because some of the observations are reserved for independent testing. Data for validating the models based on the combined data for 1986 and 1987 includes a random sample of one-half of the observations from September. These earlier observations provide a better test of the ability of the models to forecast than do the October observations.

A mean square error (MSE) suggests the degree of accuracy for a model's predictions and can be used to compare the performance of two or more models. The mean square error is computed from the forecast error e_i. The forecast error e_i is the difference between Y_i, the final ear weight for the ith sample field and $Yhat_i$, the ear weight forecast for the same field. Computation of the forecast error from independent (validation) data is conducted using coeffi-

cients estimated from the base (estimation) data to predict final ear weight. MSE can be expressed as follows:

$$MSE = (1/n) * [\Sigma (e_i)^2] \qquad (4)$$

where,

n = the number of observations in the validation data.

Comparison of MSE within years shows differences between the prediction capabilities of the models within a given period. Comparison across years shows how the models predict over time. If MSEs establish differences in the abilities of the models to forecast ear weight, then determination of differences that are statistically significant is necessary. One approach is to conduct a one-way analysis of variance (ANOVA) on the differences between the forecast errors of the models. ANOVA can establish when differences occur among the forecasts but cannot determine which pairs of forecasts are different. A multiple comparison procedure determines which forecasts are different from each other. The Tukey multiple comparison procedure is used. Appendix C shows how the ANOVA and Tukey multiple comparison procedures determine differences among the forecast errors.

The next section gives results obtained when the estimation and validation procedures discussed above were applied to data from the Corn Ear Weight Study.

RESULTS

This section presents results for estimation and validation of the corn ear weight models. Two research models were selected for comparison with the operational models. The first subsection gives estimated parameters, observed significance levels of the independent variables and other regression statistics, including R^2s and estimation MSEs. The term "estimation MSE" refers to the MSE obtained using the same data for estimation and validation. This is not the same as the MSE defined in (4), which is derived using independent data for validating the models.

Validation results reported in the second subsection show 1) if the estimated coefficients are stable over time, and 2) the level of differences in the abilities of the models to forecast the final ear weight accurately. Analyses were performed using SAS software for personal computers[4,5].

Model Performance

Several research models were estimated using the variables shown in Figure 1 (page 5). The models were similar with respect to the values of their estimation statistics. Two models whose estimation statistics were slightly better than the others' are discussed here. These models employ (S1,BT) and (V2,BD) as independent variables and are referred to as research model 1 (R1) and research model 2 (R2), respectively, in this paper. The research models are:

$$\text{Model R1:} \quad \text{Ear Weight}_i = \beta_o + \beta_1 S1_i + \beta_2 BT_i + e_i$$

$$\text{Model R2:} \quad \text{Ear Weight}_i = \beta_o + \beta_1 V2_i + \beta_2 BD_i + e_i$$

Definitions of variables included in the models are

$S1 = HL*BD$
$V2 = (HL*TD^2)/4 * \pi$
$BT = BD*TD$

where,

BD = diameter of the cob measured one inch from the butt,
TD = diameter of the cob measured two inches from the tip, and
HL = diameter of the cob measured over the husk.

Average kernel row length measured over the husk, KRL, is not included in either research model because models including this variable did not perform better than models including the husk length variable. Furthermore, the husk length is easier to measure and easier to objectively determine. The kernel row length variables should be better for estimating ear length than the cob length variable, because they reflect differences that occur in kernel row length because of factors such as the quality of pollination; however, both the research and operational kernel row length variables have flaws that cancel their potential superiority over the husk length variable. Measurement of the operational kernel row length occurs outside one of the units and is for only five ears of corn. The research kernel row length measurement takes place inside both units; however, the measurement is outside the husk, which makes determining the length of the rows more difficult.

The operational models have HL and KL as independent variables. KL is the average kernel row length measured with the husk pulled back. As stated earlier, two procedures were used to estimate the operational model. The first procedure models the relationship of KL and HL to final ear weight individually and weights the estimates together based on the R^2's of the HL and KL models. This 'weighted' model is referred to as current procedure, separate model (C_S). The second procedure jointly estimates the relationship of KL and HL to final ear weight in a single model. This model is referred to as the current procedure, joint model (C_J).

Analyses to detect the existence of outliers in the data identified some influential values; a review of these values found none to be invalid, so all values were kept.

Table 1 contains independent variables, estimated slope coefficients, P values, estimation MSEs and R^2s for the research and operational models. It shows that parameters estimated for variables in all models, research and operational, have observed significance levels equal to .0001. Estimation MSEs of the research models are smaller than those of the operational models and R^2s of the research models are almost twice as high as those of the operational models. Both these measures of model efficiency confirm the improved properties of the research models over the operational models.

Table 1. Estimation statistics for Objective Yield research and operational models constructed within and across years.

Year	Model	Independent Variables	Estimated Coefficients	P Value	Estimation MSE	R^2
1986	R1	S1	.000519	.0001		
		BT	.000104	.0001	.00199	.6092
	R2	V2	.000103	.0001		
		BD	.010303	.0001	.00205	.5975
	C_J	KL	.020516	.0001		
		HL	.030677	.0001	.00322	.3670
	C_S	KL	.024564	.0001	.00419	.1698
		HL	.034228	.0001	.00378	.2513
1987	R1	S1	.000386	.0001		
		BT	.000141	.0001	.00238	.7256
	R2	V2	.000115	.0001		
		BD	.009666	.0001	.00237	.7261
	C_J	KL	.021675	.0001		
		HL	.031384	.0001	.00476	.4512
	C_S	KL	.032526	.0001	.00586	.3138
		HL	.044579	.0001	.00568	.3396
1986 & 1987	R1	S1	.000432	.0001		
		BT	.000126	.0001	.00242	.6738
	R2	V2	.000109	.0001		
		BD	.009709	.0001	.00242	.6740
	C_J	KL	.021245	.0001		
		HL	.030124	.0001	.00428	.4230
	C_S	KL	.030663	.0001	.00525	.2884
		HL	.041547	.0001	.00510	.3079

Validation

Two methods of validation provide a means of comparing the models. The first method determines if estimated coefficients are stable and the second method assesses the accuracy of the ear weight predictions.

Stability of Coefficients

The F test given in (B1), Appendix B, determines if the estimated coefficients are stable; that is, if the equations estimated by year are statistically the same. Estimation of each model uses the data for 1986, the data for 1987 and the data for both years. Table 2 gives the estimated coefficients and the percent by which the coefficients change between years (1986 to 1987). The table shows that coefficients for model R1 change considerably from 1986 to 1987 and that the change for coefficients of model R2 is moderate.

Table 2. Coefficients estimated by year and across years for research models.

Model	Variable	Estimated Coefficients 1986	1987	1986&1987	Percent Change Between 1986 & 1987
R1	S1	.000519	.000386	.000442	34
	BT	.000104	.000141	.000012	26
R2	V2	.000103	.000115	.000112	10
	BD	.010303	.009666	.009116	7

Before performing the F test for the stability of coefficients, satisfaction of the test's assumption that the variance of the regression equation for the first year is equal to the variance of the regression equation for the second year is examined. Results

of the test for the homogeneity of variances, which is explained in Appendix B, is given in Table 3. The variances are significantly different if the value of the test statistic, F'' in (B2) is outside the interval [F1, F2]. Table 3 shows that F'' falls between F1 and F2 for both research models R1 and R2.

Table 3. Error sum of squares(SSE) and F statistics for testing equality of variances.

Model	Variables	SSE 1986	SSE 1987	Degrees of Freedom	F''	F_1	F_2
R1	(S1,BT)	.25258	.30445	(128,129)	.836	≈.699	≈ 1.430
R2	(V2,BD)	.26010	.30380	(128,129)	.863	≈.699	≈ 1.430

Since the assumption of homogeneity of the variances is satisfied, the test to assess stability of the coefficients is conducted. The coefficients are statistically the same if F', the test statistic in (B1) is greater than F. Table 4 shows that F' is greater than F for model R1, which means that one or more of the coefficients of this model are unstable. F' is less than F for model R2. Therefore, the estimated coefficients of R2 are stable over time.

Table 4. Error sum of squares(SSE) and F statistics for testing equality of estimated coefficients.

Model	Variables	1986	1987	SSE Full Model	SSE Reduced Model	Degrees of Freedom	F'	F
R1	(S1,BT)	.25258	.30445	.55703	.57322	(2, 257)	2.40	2.13
R2	(V2,BD)	.26020	.30380	.56390	.57717	(2, 257)	2.00	2.13

Accuracy in Forecasting

Computation of MSEs defined in (4) employ three types of validation data. Validation of 1986 models uses 1987 data and validation of 1987 models uses 1986 data. The third data set consists of a random sample of one-half of September's observations for both years. These observations were not used in estimating the models based on the combined data set.

Table 5 gives the MSEs of the research and operational models and shows that both research models have smaller mean square errors than the operational models for all validation data.

Table 5. Mean square errors for objective yield research and operational models.

Estimation Data	Validation Data	Model	Variables	MSE	Percent Reduction in Forecast Error R1	Percent Reduction in Forecast Error R2
1986	1987	R1	(S1,BT)	.00258		
		R2	(V2,BD)	.00257		
		C_J	(HL,KL)	.00476	45.8	46.0
		C_S	(HL,KL)	.00540	52.2	52.4
1987	1986	R1	(S1,BT)	.00219		
		R2	(V2,BD)	.00221		
		C_J	(HL,KL)	.00327	33.0	32.4
		C_S	(HL,KL)	.00328	33.2	32.6
1986 & 1987	1968 & 1987	R1	(S1,BT)	.00177		
		R2	(V2,BD)	.00183		
		C_J	(HL,KL)	.00327	45.9	44.0
		C_S	(HL,KL)	.00345	48.7	47.0

Furthernore, Table 5 shows that the reduction in MSE when forecasting with the research model is at least 32 percent. The reduction in mean square error is greatest for the 1987 data. The 1987 growing season in Michigan was drier than the 1986 growing season, which may explain the greater reduction in mean square

error for the 1987 data. Comparing the mean square errors is one method of comparing the abilities of the models to forecast. Another method is determining if differences between actual forecast errors are statistically significant.

Nonparametric ANOVA and Tukey multiple comparison procedures determine if forecast error differences are significant. The procedures which are given in Appendix C, (C1) to (C3), are applied to ranks of absolute values of the forecast errors.

Table 6 gives the nonparametric ANOVA results. For each type of validation data, two types of comparisons are made. The first comparison examines differences among forecast error means of research model R1 and operational models C_S and C_J. Likewise, the second comparison examines differences among forecast error means of research model R2 and operational models C_S and C_J. The last column of the table shows observed significance levels for each type of validation data. The observed levels are all less than .02 which leads to rejection of the null hypothesis that the forecast effect is essentially zero.

Table 6. Analysis of variance results for determining significance of differences between means of forecast errors of operational models and each research model.

Estimation Data	Validation Data	Models	Number of Obs.	MSTR	MSE	F_R	Pr > F
1986	1987	R1,C_S,C_J	393	93868	12488	7.52	.0006
		R2,C_S,C_J	393	91566	12500	7.33	.0008
1987	1986	R1,C_S,C_J	390	57244	12477	4.59	.0107
		R2,C_S,C_J	390	54348	12492	4.35	.0135
1986 & 1987	1986 & 1987	R1,C_S,C_J	228	36012	4070	8.85	.0002
		R2,C_S,C_J	228	30962	4114	7.53	.0007

Although Table 6 shows that at least two of the three models in each type of comparison have significantly different forecast errors, the table does not show which two are different. The Tukey multiple comparison procedure which constructs a simultaneous set of confidence intervals for which the experimentwise error rate is .10, provides a means of determining when pairs are statistically different. Two types of forecast errors are essentially the same if a confidence interval of the difference between their means contains zero. Table 7 gives simultaneous upper and lower confidence levels constructed for the validation data sets. The table shows that differences between forecast errors of the research and operational models are significant.

Table 7. Tukey's studentized range test for differences between ranks of forecast errors (*** Comparison significant at the .10 level).

Estimation Data	Validation Data	Treatment Comparison	Simultaneous Upper and Lower Confidence Limits	
1986	1987	R1-C_J	(-60.74 , -3.90)	***
		R1-C_S	(-81.54 ,-24.70)	***
		C_J-C_S	(-49.22 , 7.62)	
		R2-C_J	(-60.06 , -3.19)	***
		R2-C_S	(-80.95 ,-24.08)	***
		C_J-C_S	(-49.32 , 7.55)	
1987	1986	R1-C_J	(-65.38 , -8.34)	***
		R1-C_S	(-64.33 , -7.29)	***
		C_J-C_S	(-27.46 , 29.57)	
		R2-C_J	(-64.20 , -7.13)	***
		R2-C_S	(-63.69 , -6.62)	***
		C_J-C_S	(-28.02 , 29.05)	
1986 & 1987	1986 & 1987	R1-C_J	(-55.20 , -12.51)	***
		R1-C_S	(-61.98 , -19.29)	***
		C_J-C_S	(-28.12 , 14.57)	
		R2-C_J	(-52.75 , -9.82)	***
		R2-C_S	(-59.20 , -16.27)	***
		C_J-C_S	(-29.92 , 15.01)	

CONCLUSIONS

Data from the 1986 and the 1987 Michigan Corn Ear Weight Study were used to estimate models employing surface area or volume measurements as independent variables. These models had mean square errors that are 32 to 52 percent lower than those for operational models for the 1986 and 1987 validation data. Mean square errors of the research models estimated using data for both years were 44 to 48 percent lower than those of the operational models for the same period. Tests of statistical hypotheses and multiple comparison procedures conducted suggested that differences between forecast errors of the research and operational models were significant.

Results from analyses to determine if the research models have stable coefficients showed that one of the models had coefficients that were unstable. This model had S1 (surface area of the corn ear) and BT(the interaction between the butt and tip diameters of the corn ear) as independent variables. The second research model had stable coefficients across years. It employed butt diameter and volume of the corn ear as independent variables. Analysis of both models will continue when data for the 1988 Corn Research Project is available. Recommendations regarding the research models will be given after analyses of the 1988 data are completed.

1988 CORN EAR WEIGHT STUDY

The Corn Ear Weight Study continued in 1988. The features of the 1988 project are discussed below.

1. The project continued in Michigan in 1988 and was also expanded to Missouri to evaluate the performance of the surface area and volume models in a corn environment that is more drought prone than Michigan. In retrospect, both states suffered drought conditions in 1988. The analysis of the 1988 data should provide valuable information about the research models.

2. Because of the difficulty in determining the average length of the kernel row by feeling through the husk, the cob length measurement from the operational program replaced the kernel row length measurement over the husk (KRL).

Once it was established that the sample field met the requirement of being in maturity category 4, 5, or 6, the enumerator no longer was asked to determine if each ear in the unit satisfied this requirement for inclusion in the study. The enumerator measured all ears in row 1 of each unit. This eliminated some subjectivity in the study, especially the difficulty in determining level of maturity of an ear of corn without removing the husk.

3. The project has been expanded to include the estimation of ear weight at both the sample level and the individual ear level. This was done by instructing laboratory personnel to weigh the 3rd and 4th ears from row 1 of each unit separately to obtain their individual weights. Individual ear weight models will use these weights for estimation. The forecasts produced by the two methods will be compared.

4. Data from the 1988 study will determine the optimum number of units to sample and the minimum number of ears to measure in estimating ear weight accurately.

REFERENCES

1. Belsley, David A., Kuh, Edwin and Roy E. Welsch. Regression Diagnostics: Identifying Influential Data and and Sources of Collinearity, New York: John Wiley & Sons, 1980

2. Conover, W.J., and Ronald L. Mann. Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics, The American Statistician, August 1981, Vol. 35, No. 3.

3. Neter, John and William Wasserman, Applied Linear Statistical Models, Illinois: Richard D. Irwin, Inc., 1977.

4. SAS Institute Inc. SAS Procedures Guide for Personal Computers, Version 6 Edition. Cary, NC: SAS Institute inc., 1985

5. SAS Institute Inc. SAS/STAT Guide for Personal Computers, Version 6 Edition. Cary, NC: SAS Institute inc., 1985

6. Steele, Ronald J., 1985. Preliminary Investigation For An Improved Corn Grain Weight Forecast Model, Unpublished.

7. U.S. Department of Agriculture, 1987 Corn Objective Yield Survey: Enumerator's Manual, National Agricultural Statistics Service, Washington, D.C., 1987

8. U.S. Department of Agriculture, 1987 Corn Objective Yield Survey: Supervising & Editing Manual, National Agricultural Statistics Service, Washington, D.C., 1987

9. U.S. Department of Agriculture, 1987 Michigan Corn Research Study, National Agricultural Statistic Service, Washington, D.C., 1987

10. U.S. Department of Agriculture, Statistical Reporting Service, Scope and Methods of the Statistical Reporting Service, Washington, D.C., 1983

11. Vogel, Frederic A., Using Corn Plant Vegetative Characteristics To Forecast Production of Grain Per Plant, U.S. Department of Agriculture, Statistical Reporting Service, Washington, D.C., 1971

Appendix A. Detection of Outliers

$DFFITS_i$, a statistic proposed by Belsley, Kuh and Welch[1] is used for outlier detection in this study. $DFFITS_i$ measures the influence of the ith observation in predicting the value of Y_i (ear $weight_i$).

$$DFFITS_i = X_i[b - b_i(i)/[s(i)* (X_i(X'X)^{-1}X'_i)^{\frac{1}{2}}]$$

$$= Y_i - Y_i(i)/[s(i)* (X_i(X'X)^{-1}X'_i)^{\frac{1}{2}}]$$

where,

b = the estimated regression coefficient using all of the data in X,
$b(i)$ = the estimated regression coefficient when X_i is deleted,
Y_i = the predicted value when the ith observation is included in the estimation data,
$Y_i(i)$ = the predicted value when the ith observation is excluded, and
$s(i)$ = variance of the error term when the ith observation is deleted.

Outliers are identified by absolute values of DFFITS greater than $2*[(p/n)^{\frac{1}{2}}]$

where,

p = the number of parameters estimated, and
n = the number of observations.

Appendix B. Stability of Coefficients

In this research, stability of estimated coefficients over time is analyzed by estimating the coefficients by year and using a statistical test to determine if coefficients estimated for the first year are the same as coefficients estimated for the second year [3]. The test compares the error sum of squares when the coefficients are assumed to be the same with the error sum of squares when the coefficients are assumed to be different. The larger the difference between the two sums, the higher the probability of rejecting the hypothesis that the coefficients are the same. In testing the equality of coefficients in two regression equations, each of which includes two independent variables, the hypothesis tested is as follows:

H_o: $\beta_{o1} = \beta_{o2}$, $\beta_{11} = \beta_{12}$, $\beta_{21} = \beta_{22}$

H_a: $(\beta_{o1} \neq \beta_{o2})$ and/or $(\beta_{11} \neq \beta_{12})$ and/or $(\beta_{21} \neq \beta_{22})$

where,

β_{o1}, β_{11}, β_{21} = coefficients estimated for the first regression equation, and

β_{o2}, β_{12}, β_{22} = coefficients estimated for the second regression equation.

The null hypothesis is rejected if the test statistic

$$F' = \frac{[SSE(R) - SSE(F)]/p}{SSE(F)/(n1+n2-2p)} \tag{B1}$$

is greater than $F(1-\alpha;p,n1+n2-2p)$, the $1-\alpha$ percentile of the F distribution with p and n1+n2-2p degrees of freedom. The number of parameters estimated in the regression function is p. The probability of rejecting the null hypothesis given the hypothesis is true is α (.10). In (B1), SSE(R) is the error sum of squares obtained by assuming the coefficients are equal (model called reduced) and estimating the equation for both years. SSE(F) is the error sum of squares obtained by assuming the coefficients are different (model called full). Computation of SSE(F) involves estimating the equations for the first and second years separately and adding the individual sum of squares. The number of observations for the first and second years are n1 and n2, respectively.

The test for determining stability of the coefficients is based on the assumption that variances of the error terms from the regression for the first year (SSE_1) and second year (SSE_2) are equal. Performing the F test for the equality of variances [3] can determine the validity of this assumption. The test leads to rejection of the null hypothesis of equal variances if

$$F'' = \{(SSE_1)/(n1-p)\} / \{(SSE_2)/(n2-p)\} \qquad \text{(B2)}$$

is greater than F1 or less than F2. F1 is the $\alpha/2$ percentile of the F distribution with (n1-p, n2-p) degrees of freedom and F2 is the $1-\alpha/2$ percentile of the F distribution with (n1-p, n2-p) degrees of freedom with α equal to .05. When rejecting the null hypothesis, the test for determining if the coefficients are stable may be invalid, unless a transformation of the data equalizes the variances.

APPENDIX C. One-way Analysis of Variance and Tukey Multiple Comparison Procedure for Differences Among Forecast Errors

When mean square errors show differences in the ability of models to forecast ear weight, then other analyses can determine if the differences are statistically significant. One approach would be to conduct a one-way analysis of variance (ANOVA) on the differences between the forecast errors for the models. A one-way ANOVA defines the following model:

$$Y_{ij} = \mu. + \tau + e_{ij} \qquad (C1)$$

where,

Y_{ij} = the absolute value of the forecast error for the ith sample field and jth model,
$\mu.$ = the overall mean,
τ = the effect of of the forecast error for the jth model, and
e_{ij} = the deviation of $(\mu. + \tau)$ from Y_{ij}.

The analysis of variance uses the model in (C1) to test the following hypotheses:

H_o: $\tau_1 = \tau_2 = \tau_3 = \ldots \tau_k$, k = the number of models
H_a: not all τ_j's are zero, j = 1,2,...,k

The test uses a nonparametric procedure [2] since the statistical distribution of the absolute values of the forecast errors is unkown. The test statistic used is as follows:

$$F_R = \frac{[\ \Sigma\{(R^2{}_j/n_j) - (n(n+1)^2/4)\}]/k-1}{\Sigma\ \Sigma\ R(X_{ij}) - R_{.j})^2/n-k} \qquad (C2)$$

where,

$R(X_{ij})$ = rank of absolute value of the ith forecast error for the jth model when forecast errors for all models are ranked jointly from 1 to n,

R_j = sum of the ranks for model j's forecasts,

$R_{.j}$ = average rank for model j's forecasts,

n_j = the number of observations for model j, and
n = the total number of observations.

The numerator in F_R represents differences between ranks of the k models and the denominator represents differences within ranks of the kth model. If the null hypothesis is true and the forecast effect is essentially zero, then F_R will be small. The test rejects H_0 for values of F_R larger than the $(1-\alpha;\ k-1, n-k)$ percentile of the F distribution. The observed significance level of the test is the probability of getting a larger value of F_R than that computed for the data. The null hypothesis is rejected for observed significance levels (P values) smaller than or equal to 0.10.

Rejection of the null hypothesis implies that the forecasts are significantly different. The Tukey multiple comparison procedure determines which forecasts are different from each other and controls the experimentwise error rate at .10, in this study. The procedure is performed on ranks of absolute values of the forecast errors. Confidence intervals are constructed simultaneously for all pairs of forecasts. The interval

$$D - Ts(D) \leq \mu_j - \mu_j' \leq D + Ts(D) \qquad \text{(C3)}$$

is used to determine the difference between errors of the jth and jth' forecasts. D is the difference between average ranks of the errors, s is the standard deviation of D, and T is the $(1-\alpha,\ k,\ n-k)$ percentile of the studentized range distribution. T is defined as as follows:

$$T = (1/\sqrt{2}) * q(1-\alpha;\ k,\ n-k)$$

where,

q = the studentized range and is equal to the $\{\max(\text{of } Y_j) - \min(\text{of } Y_j)\}/\sigma$,
σ = the estimated standard deviation of k random variables with mean μ and variance σ^2 and,
n = the number of observations and k is the number of forecasts compared.

Only if the interval constructed in (C3) contains zero will the forecast errors be statistically the same. Otherwise the forecast errors are different.

UNITED STATES DEPARTMENT OF AGRICULTURE
NATIONAL AGRICULTURAL STATISTICS SERVICE

Form Approved
O.M.B. Number 0535-0088
Expiration Date 7/31/89
C.E. 12-0032B-1

FORM B—1: CORN YIELD COUNTS — August 1, 1987

YEAR, CROP, FORM, MONTH (1—4)
7431

Has operator applied pesticides with organophosphorous content since last field visit? YES ☐ NO ☐

If YES, enter latest application date ________________ *and name of pesticide* ________________

UNIT LOCATION	UNIT 1	UNIT 2
Number of rows along edge of field		
Number of paces into field		

Date (________________)	370
Starting Time *(Military Time)*	371

UNIT LOCATION CODE

1. a. First visit to lay out unit 1
   b. Unit relocated this month 2
   c. Sample unit laid out previously 3
   } Enter Code

UNIT 1	UNIT 2
302	307

Skip To Item 3 if code 3

ROW SPACE MEASUREMENTS

	UNIT 1	UNIT 2
2. a. Measure distance from stalks in Row 1 to stalks in Row 2 . . . . Feet & Tenths	303 •	304 •
b. Measure distance from stalks in Row 1 to stalks in Row 5 . . . . Feet & Tenths	305 •	306 •

COUNTS WITHIN 15-FOOT UNITS

	ROW 1	ROW 2	ROW 1	ROW 2
3. Number of stalks	331	332	333	334
4. Number of stalks with ears or silked ear shoots *(Item 4 cannot exceed Item 3 for any row)*	341	342	343	344
5. Number of ears and silked ear shoots *(Item 5 must equal or exceed Item 4 for any row)*	351	352	353	354
6. Number of ears with evidence of kernel formation *(Item 6 cannot exceed Item 5 for any row)*	361	362	363	364

OBSERVATIONS BEYOND UNIT 2, ROW 1 ONLY:

Husk the first 5 ears or silked ear shoots beyond Row 1 and examine for maturity. If ears or silked ear shoots are not yet present CHECK (   ) and skip Items 7—13.

Maturity Stage	Code	Maturity Stage	Code
Pre-Blister	2	Dough	5
Blister	3	Dent	6
Milk	4	Mature	7

	Ear Number					Total of 5 Ears
	1	2	3	4	5	
7. Maturity stage of first 5 ears or silked ear shoots						301

If total in Item 7 is
- → 12 or less, skip Items 8 through 13.
- → 13 or more, continue. (If any ears in Item 7 are Code 2, replace each Code 2 ear with the next Code 3 ear or higher and enter in Item 8.)

FORM B—1: CORN (Cont'd)

	Ear Number				
	1	2	3	4	5
8. Maturity stage of first 5 ears Code 3 or higher .... Code	320	321	322	323	324

Does Item 8 have 3 or more Code 7 ears?
- ☐ YES, *Complete Items 12 and 13 only.*
- ☐ NO, *Continue.*

9. Is the field to be harvested as High Moisture Corn within the next 3 days and were 3 or more code 6's recorded in Item 8?
   - ☐ YES, *Enter code 1 and complete Items 12 and 13 only.*
   - ☐ NO, *Enter code 2 and complete Items 10 and 11 only.*

   325

10. Average length of kernel rows (Item 8 ears) . . Inches & Tenths	326 .	327 .	328 .	329 .	330 .

MEASUREMENTS WITHIN UNIT 2, ROW 1

11. Enter the number of ears counted in Item 6 for Unit 2, Row 1 (Cell 363) ____________ .
    Measure length of cob for all ears counted. The number measured and entered below must equal the number counted in Cell 363.

    When measuring do NOT *remove "ear" or pull back husk. Record to nearest 1/10 inch. If more than 30 ears, use blank space on right.*

1. ____ . ____	6. ____ . ____	11. ____ . ____	16. ____ . ____	21. ____ . ____	26. ____ . ____	29. ____ . ____
2. ____ . ____	7. ____ . ____	12. ____ . ____	17. ____ . ____	22. ____ . ____	27. ____ . ____	30. ____ . ____
3. ____ . ____	8. ____ . ____	13. ____ . ____	18. ____ . ____	23. ____ . ____	28. ____ . ____	
4. ____ . ____	9. ____ . ____	14. ____ . ____	19. ____ . ____	24. ____ . ____	Total Length	308 .
5. ____ . ____	10. ____ . ____	15. ____ . ____	20. ____ . ____	25. ____ . ____	Total Ears	309

HARVESTING SAMPLE UNITS

	Unit 1 Row 1	Unit 2 Row 1
12. *HUSK and TAG 3rd and 4th ears in Row 1 of each unit. Then husk remaining ears.* Number of ears husked with grain (include 3rd & 4th ears) .... VERIFY: Cells 312 = 361 and 313 = 363	312	313
13. Weight of ears with grain from row 1 of each unit (include 3rd & 4th ears, exclude weight of containers) .... Pounds & Tenths	314 .	315 .

> *Place 3rd and 4th ears of Row 1 in separate plastic bags for each unit. After completing Items 12 and 13, send FORM B to the State office and send 3rd and 4th ears to the Regional Lab.*

Ending Time *(Military Time)*	372
Status Code	380
Enumerator Number	390
Supervisor Number	391

Enumerator ______________________________

Did your supervisor assist you in working this sample? Yes ☐ No ☐

UNITED STATES DEPARTMENT OF AGRICULTURE
NATIONAL AGRICULTURAL STATISTICS SERVICE

Form Approved
O.M.B. Number 0535-0088
Expiration Date 7/31/89
C.E. 12-0032C-2

FORM C—2: 1987 CORN FINAL PRE—HARVEST LAB DETERMINATIONS

MONTH CODE

Sept. 1.........2
Oct. 1.........3
Nov. 1.........4
After Nov. 1....5

YEAR, CROP, FORM, MONTH (1—4)
745 __

EAR WEIGHT (BOTH UNITS COMBINED)

1. Weight of ears in sealed bags .... Grams to Tenths [501]
2. Weight of same number of new bags and rubber bands .... Grams to Tenths [502]

GRAIN WEIGHT AND MOISTURE DETERMINATIONS

Shell grain from all ears. If ears are too wet to shell easily dry them for a short period at no more than 70 °C before shelling.

3. Weight of all grain shelled from ears at time of moisture test .... Grams to Tenths [507]
4. Moisture content of shelled grain .... Percent (One Decimal) [508]

Lab Technicians ______________________ Date Analysis Completed ______________
Month/Day

Over Please

UNITED STATES DEPARTMENT OF AGRICULTURE
NATIONAL AGRICULTURAL STATISTICS SERVICE

Form Approved
O.M.B. Number 0535-0088
Expiration Date 7/31/89
C.E. 12-00328-R

FORM B—R CORN RESEARCH STUDY

YEAR, CROP, FORM, MONTH (1—4)
743 __

UNIT 1

1. Does Item 8, Form B, have 3 or more code 4 ears but less than 3 code 7 ears?
   - ☐ YES - *Complete Item 2.*
   - ☐ NO - *Return to Form B.*

UNIT 2

1. Does Item 8, Form B, have 3 or more code 4 ears but less than 3 code 7 ears?
   - ☐ YES - *Complete Item 2.*
   - ☐ NO - *Return to Form B.*

MEASURMENTS

2. When measuring, do **NOT** remove ear or pull back husk or damage ear.

Measure only those ears that are maturity code 4 or higher.

Measure:
The average kernel row length over the husk. Record to nearest 1/10 inch.
The diameter of the ear one inch from the butt of the cob. Record to nearest millimeter.
The diameter of the ear two inches from the tip of the cob. Record to nearest millimeter.

UNIT 1 ROW 1

U	Ear No.	Kernel Row Length (to 1/10 inch)	Butt Diameter (to 1 mm)	Tip Diameter (to 1 mm)
1	1.	.		
1	2.	.		
1	3.	.		
1	4.	.		
1	5.	.		
1	6.	.		
1	7.	.		
1	8.	.		
1	9.	.		
1	10.	.		

UNIT 2 ROW 1

U	Ear No.	Kernel Row Length (to 1/10 inch)	Butt Diameter (to 1 mm)	Tip Diameter (to 1 mm)
2	1.	.		
2	2.	.		
2	3.	.		
2	4.	.		
2	5.	.		
2	6.	.		
2	7.	.		
2	8.	.		
2	9.	.		
2	10.	.		

If more than 10 ears, continue on back.